ORIGINAL PAPER

Molecular Cloning and Expression Analysis of Triticale Phytocystatins During Development and Germination of Seeds

Joanna Szewińska · Edyta Zdunek-Zastocka · Mirosław Pojmaj · Wiesław Bielawski

Abstract Three triticale cDNAs encoding inhibitors of cysteine endopeptidases, belonging to phytocystatins, have been identified and designated as *TrcC-1*, *TrcC-4* and *TrcC-5*. Full-length cDNAs of *TrcC-1* (617 bp) and *TrcC-4* (940 bp), as well as a fragment of *TrcC-5* cDNA (369 bp), were obtained. A high-level identity of the deduced amino acid sequence of TrcCs with other known phytocystatins, especially with wheat and barley, has been observed. Moreover, the presence of conserved domain, containing the G and W residues, the sequence of QxVxG and the sequence of LARFAV, characteristic for plant cysteine endopeptidase inhibitors, has been noted. The profiles of *TrcC-1* and *TrcC-5* mRNA levels in the developing seeds of two triticale cultivars that differ in their resistance to preharvest sprouting (Zorro and Disco) were similar. However, the expression of *TrcC-4* was, higher in the developing seeds, and in the scutellum of germinating seeds of a cultivar more resistant to preharvest sprouting (Zorro) than in the less resistant (Disco). Additionally, the expression of *TrcC-4* remained longer in developing seeds of Zorro as compared to Disco. The performed studies suggest that TrcC-4 might have an influence on the higher resistance of Zorro cultivar to preharvest sprouting.

J. Szewińska (✉) · E. Zdunek-Zastocka · W. Bielawski
Department of Biochemistry,
Warsaw University of Life Sciences-SGGW,
Nowoursynowska 159,
02-776 Warsaw, Poland
e-mail: joanna_szewinska@sggw.pl

M. Pojmaj
DANKO Plant Breeders Ltd.,
Laski,
05-660 Warka, Poland

Introduction

Cysteine endopeptidases are enzymes that possess, in an active centre, a reactive cysteine residue crucial for hydrolysis of peptide bonds in plant and animal proteins. Despite important role of cysteine endopeptidases, their uncontrolled activity can induce pathology, as exemplified by the secretion of lysosomal proteases and the destruction of tissue structures by extracellular peptidases of bacterial pathogens for invasion of new tissues, as well as the undesirable phenomenon in cereal grains to undergo preharvest sprouting (Travis et al. 1995). To prevent such threats, many organisms have evolved protective mechanisms that are activated at the level of transcription, translation and post-translational modifications. The most direct regulators of proteinases that are synthesised by the cell are protein inhibitors (Bode and Huber 1992; Viswanathan et al. 2011). Inhibitors that regulate the activity of cysteine proteinases are called cystatins. In animals, there are three types of cystatins: stefins, cystatins and kininogens. In plants, the cystatins are called phytocystatins. Phytocystatins (PhyCys) share with animal cystatins three regions that are involved in the interaction with their target enzymes: the sequence of QxVxG, a glycine residue located near the N-terminus and a tryptophan residue located near the C-terminus. The presence of an additional Leu-Ala-Arg-Phe-Ala-Val (LARFAV) sequence, which is characteristic for PhyCys, probably allows for regulation of the cysteine endopeptidases from the papain family (Abe et al. 1987; Gburek and Gołąb 1996; Margis et al. 1998; Volpicella et al. 2011). Most phytocystatins are small proteins, ranging in size from 12 to

16 kDa, with a distinct group having a higher molecular weight (23 kDa), due to a C-terminal extension (Shyu et al. 2004). However, in potato and tomato, multi-cystatin proteins with molecular weights of approximately 85 kDa (Walsh and Strickland 1993; Wu and Haard 2000) were found as well.

The location of cystatins in the tissues of living organisms is often related to their function. Phytocystatins that regulate gene expression in vegetative and generative organs protect those parts of plants against the exogenous endopeptidases from pests (Zhang et al. 2011) and protect the healthy tissues from excessive proteolysis from endogenous endopeptidases (Gburek and Gołąb 1996; Muellenborn et al. 2011). Experiments conducted on transgenic plants have confirmed the role of proteinase inhibitors in the fight against pests (Carrillo et al. 2011). It has been observed that the grey garden slug *Deroceras reticulatum* develops slower when it feeds on *Arabidopsis* leaves overexpressing a rice cystatin gene (OC-I) (Mosolov and Valueva 2005). Cystatin gene expression is also upregulated in response to adverse environmental conditions (Prabu et al. 2011), such as oxygen deficiency, poor lighting or low temperatures (Gaddour et al. 2001). The previous studies have also suggested that phytocystatins in tissues of developing and germinating seeds participate in the regulation of peptidase activity, which takes part in the mobilisation of storage proteins. It has been shown that proteins of phytocystatins from corn (CC), barley (Hv-CPI) and wheat (WC) are present in embryonic cells and in the aleuron layer, where *de novo* synthesis of cysteine endopeptidases occurs (Abe et al. 1987, 1994; Gaddour et al. 2001; Kuroda et al. 2001). Phylogenetic analysis performed by Martinez et al. (2009) compared 13 barley phytocystatins (HvCPI-1 to HVCPI-13) with the available amino acid sequences of other phytocystatins from the Poaceae family, and their results led to the division of these inhibitors on three functional groups, namely, A, B and C, and two subgroups, namely, C1 and C2. The results indicated that the barley phytocystatins (HvCPI-1–4), as well as other inhibitors that belong to group A, were present in generative and vegetative tissues. Therefore, it was proposed that these cystatins might be universal in nature. Members from the B and C cystatin groups demonstrated a preference towards the enzymes of seed tissues (Abraham et al. 2006; Martinez et al. 2009). Despite the performed studies (Martinez et al. 2009), the role of phytocystatins in cereal seeds has not been completely explained. The changes in the global cereal market tend to try to clarify the functions of these inhibitors in cereal species such as triticale. Triticale seeds in favourable wetting, aeration and temperature conditions are particularly susceptible to premature germination in the ear, which is termed preharvest sprouting. Among the triticale cultivars, most are not completely resistant to preharvest sprouting; however, there are some cultivars that

are more resistant than others. Our previous studies showed that cysteine endopeptidases play a unique role in a process of preharvest sprouting by activation of a premature degradation of seed storage proteins (Bielawski et al. 1994; Bielawski and Prabucka 2001; Prabucka and Bielawski 2004; Drzymała et al. 2009). Therefore, it is vital to determine whether the presence of phytocystatins in developing and germinating triticale seeds determines the level of resistance of the cultivar to premature germination.

The aim of the presented research was to identify and characterise the nucleotide sequences of phytocystatins in developing and germinating triticale seeds. In this study, for the first time, the mRNA levels of phytocystatins in developing and germinating seeds of two cereal cultivars, extreme in terms of their resistance to preharvest sprouting, were determined as well. The possible role of the identified inhibitors in the control of cysteine endopeptidases activity co-responsible for the process of preharvest sprouting in the ear is discussed.

Materials and Methods

Plant Material

The seeds of two cultivars of winter triticale (Zorro and Disco), which differ greatly in their resistance to preharvest sprouting, were obtained from Danko Plant Breeders Ltd. in Laski, Poland. The Zorro cultivar is known as one of the most resistant cultivars to preharvest sprouting, while the Disco cultivar is considered to be one of the most sensitive cultivars to preharvest sprouting. The developing seeds were collected 7, 10, 13, 17, 20, 27, 34, 41 and 48 days after pollination (dap). Dry seeds underwent the imbibition process and were germinated at 22°C. The seeds were collected after 2, 6, 12, 24, 48, 72 and 96 h from imbibition, and the scutellum was separated from the endosperm. The prepared material was frozen in liquid nitrogen and stored at −80°C until further use. The shoots and roots of the developing seedlings were removed and were not the subject of this study.

Isolation of the Total RNA, cDNA Synthesis and Amplification of the Internal Regions of the *TrcCs*

The total RNA was isolated from the developing and germinating seeds following the method of Chomczynski and Sacchi (Chomczynski and Sacchi 1987), with initial extraction performed in a buffer containing 50-mM Tris–HCl pH 9.0, 200-mM NaCl, 1% sarcosyl, 20-mM EDTA and 5-mM DTT and further extraction in a mixture of phenol/chloroform/isoamyl alcohol. Based on the total RNA, the cDNA templates were synthesised using oligo

(dT) primers according to manufacturer's protocol (Promega Corp., Madison, USA). Internal fragments of phytocystatins (TrcCs) were amplified using the following specific primers listed in Table 1. The primers were designed based on wheat, barley, rice and corn phytocystatin sequences, which are available in the NCBI database (http://www.ncbi.nlm.nih.gov). PCR conditions were as follows: 3 min at 94°C; 35 cycles of 30 s at 94°C, 30 s at 52°C (TrcC-1) or 53°C (TrcC-4) or 54°C (TrcC-5) and 2 min at 72°C; and a final 7 min at 72°C. The amplification products were ligated into pGEM-T Easy vector and amplified in E. coli JM109 cells (Promega Corp., Madison, USA). The sequencing was performed at the Institute of Biochemistry and Biophysics, Polish Academy of Sciences.

5′ and 3′ Rapid Amplification of cDNA Ends of TrcC-1 and TrcC-4 (5′ and 3′ RACE)

To obtain the full-length cDNA of TrcC-1 and TrcC-4, the GeneRacer Kit (Invitrogen Corporation, Carlsbad) was used. The 5′ and 3′ ends of TrcC-1 and TrcC-4 were amplified in two PCR stages with the use of primers that were specific to the selected genes (Table 1) and with the GeneRacer universal primers (included in the Invitrogen kit).

First round of PCR for the 5′ end amplification of TrcC-1 and TrcC-4 was performed under the following conditions: 2 min at 94°C; 5 cycles of 30 s at 94°C and 1 min at 72°C; 5 cycles of 30 s at 94°C and 1 min at 70°C; 25 cycles of 30 s at 94°C, 30 s at 62°C, 1 min at 68°C and a final 10 min at 68°C. The PCR product of the first round of PCR was diluted ten-fold and was used as the template in the second PCR stage (nested PCR). The nested PCR for amplification of the 5′ end of TrcC-1 and TrcC-4 was performed using the following conditions: 2 min at 94°C; 35 cycles of 30 s at 94°C,

30 s at 62°C (TrcC-1) or 61°C (TrcC-4), 30 s at 68°C and a final 10 min at 68°C.

First round of PCR for the 3′ end amplification of TrcC-1 and TrcC-4 was performed under the following conditions: 2 min at 94°C; 5 cycles of 30 s at 94°C and 1 min at 72°C; 5 cycles of 30 s at 94°C and 1 min at 70°C; 25 cycles of 30 s at 94°C, 30 s at 65°C, 1 min at 68°C and a final 10 min at 68°C. The PCR product of the first round of PCR was diluted ten-fold and was used as the template in the second PCR stage (nested PCR). The nested PCR for the amplification of the 3′ end of TrcC-1 and TrcC-4 was performed using the following conditions: 2 min at 94°C; 30 cycles of 30 s at 94°C, 30 s at 65°C (TrcC-1, TrcC-4), 30 s at 68°C and a final 10 min at 68°C.

Relative Quantitative RT-PCR

The mRNA levels of the three triticale phytocystatins (TrcC-1, TrcC-4 and TrcC-5) were analysed at 9 stages of seed development, after 7, 10, 13, 17, 20, 27, 34, 41 and 48 days from pollination using the Titanium One-Step RT-PCR Kit (Clontech Laboratories, USA). The mRNA levels of the phytocystatin in the germinating triticale seeds were measured after 2, 6, 12, 24, 48, 72 and 96 h from imbibition. The oligonucleotide primer sequences used in the reverse transcriptase-polymerase chain reaction (RT-PCR) reactions are shown in Table 2. To ensure that equal amounts of the RNA templates were added to each reaction, amplification of the 18S rRNA gene was performed. RT-PCR conditions for the amplification of particular phytocystatins and 18S rRNA gene were as follows: 60 min at 50°C and 5 min at 94°C; 28 cycles (TrcC-1) or 30 cycles (TrcC-4, TrcC-5) or 22 cycles (18S rRNA) of 30 s at 94°C, 30 s at 64°C (TrcC-1) or 55°C (TrcC-4) or 65°C (TrcC-5) or 55°C (18S rRNA), 45 s at 68°C and a final 2 min at 68°C. The RT-PCR generated fragments were cloned into the

Table 1 Primers used for cloning the full-length TrcC cDNAs from triticale

Primer	Oligonucleotide sequence (5′-3′)	Application	Position in cDNA (bp)	Source of primers
1-F	5′GATCGATCGATGGAGATGTG-3′	1st PCR for TrcC-1	1	WC1, HvCPI-2,
1-R	5′GAGAACACGCTAAACCATG-3′	1st PCR for TrcC-1	505	OC-I, CC1
4-F	5′CCTCGTCCTACCGTCTATTC-3′	1st PCR for TrcC-4	289	WC4, HvCPI-2,
4-R	5′CTCGCAAGAAGCAGACTTG-3′	1st PCR for TrcC-4	767	OC-I, CC1
5-F	5′GCTTCCTCCTCATCATCGTTG-3′	1st PCR for TrcC-5	1	WC5
5-R	5′CTGGCGTTTGTTAGTCCACTTCTG-3′	1st PCR for TrcC-5	348	
F-Trc	5′CTCGCCCGCTTCGCCGTCTCCGAG-3′	3′ RACE for TrcC-1, -4	217,428	TrcC-1, TrcC-4
R-Trc	5′CATAGAGCTTCTTTGCCCCGCCTTC-3′	5′ RACE for TrcC-1, -4	344,554	
F-TrcC1	5′GGTGTGGGAGAAACCATGGGAGAAC-3′	3′ RACE (nested)TrcC-1	376	TrcC-1
R-TrcC1	5′TCGAGGTCGTTCTCCTGGCCCATC-3′	5′ RACE (nested)TrcC-1	184	
F-TrcC4	5′CTGGAGCAGCTGTGGCTCGATGTC-3′	3′ RACE (nested)TrcC-4	590	TrcC-4
R-TrcC4	5′CGACGATGTAGGGGTCCTTCTCTC-3′	5′ RACE (nested)TrcC-4	402	

F forward primer, R reverse primer

Table 2 Primer sequences that were used for analysis by relative quantitative RT-PCR

Gene	Oligonucleotide sequence (5′-3′)	Productlength (bp)	Location in cDNA (bp)
TrcC-1	F: 5′-GCAAGAGAACGACCTCGAC-3′	267	177
	R: 5′-ATGAGTTATGCGCTTGGAAC-3′		424
TrcC-4	F: 5′-CAGACCCATTCCCAGAAG-3′	426	307
	R: 5′-TGCATGTCGCACGGATGG-3′		715
TrcC-5	F: 5′-GCTTCCTCCTCATCATCGTTG-3′	368	1
	R:5′CTGGCGTTTGTTAGTCCACTTCTG-3′		348
18S rRNA	F: 5′-GATCCATTGGAGGGCAAGTC-3′	244	27
	R: 5′-GATGGCTTGCTTTGAGCACTC-3′		224

pGEM-T Easy vector, and at least six clones of each insert were sequenced and analysed. Three repetitions of RT-PCR amplification, based on the cDNA that was obtained from two independent RNA extractions, were performed.

Computer Analysis of the Nucleotide and Amino Acid Sequences

Sequence homology of the triticale phytocystatins was verified by database searches at NCBI (http://www.ncbi.nlm.nih.gov/) with the use of the BLAST algorithm (http://blast.ncbi.nlm.nih.gov/). The deduction of the amino acid sequences and the calculation of the theoretical molecular weight and the isoelectric point (pI) were performed using the ExPASy Proteomics tools (http://www.expasy.ch/tools/). The percentage identity of the amino acid sequences was calculated using ClustalW software (http://www.clustal.org/).

ClustalW multiple protein sequence alignment was used to construct a phylogenetic tree of the phytocystatins of the Poaceae family using the maximum likelihood method. The maximum likelihood tree was built using PhyML (Dereeper et al. 2008) with the JTT (Jones-Taylor-Thornton) amino acid substitution model and 100 bootstrap replicates.

Results

Nucleotide Sequence Analysis of cDNA of TrcCs

Based on the cDNA templates that were obtained from the triticale seeds of the Zorro and Disco cultivars, three cDNAs encoding phytocystatins were amplified and designated as TrcC-1, TrcC-4 and TrcC-5. The introduced numbering of the triticale phytocystatins corresponded to the numbering of wheat phytocystatins (WC1–WC5; Kuroda et al. 2001; Corre-Menguy et al. 2002) to which the obtained triticale nucleotide sequences revealed the highest percentage of identity. The lengths of the TrcCs that were obtained in the first PCR were as follows: 267 bp for TrcC-1, 426 bp for

TrcC-4 and 369 bp for TrcC-5. Afterwards, the TrcC-1 and TrcC-4 fragments were elongated in the direction of the 5′ and 3′ ends (5′ and 3′ RACE method). Consequently, for TrcC-1, a 617-bp cDNA was obtained, which contains the open reading frame (ORF) of 429 bp, the 5′ untranslated region (5′ UTR) of 9 bp and the 3′ UTR of 179 bp. The full length of TrcC-4 cDNA is 940 bp in length and contains a 444-bp ORF, a 5′ UTR of 201 bp and a 3′ UTR of 295 bp. The TrcC-1, TrcC-4 and TrcC-5 sequences were registered in GenBank at NCBI with the following accession numbers: TrcC-1, GU395200; TrcC-4, GU395201; and TrcC-5, GU395202. The cDNA of TrcC-1 encodes a protein of 142 amino acids with a calculated molecular weight of 15.7 kDa and a pI of 5.73, whereas the TrcC-4 cDNA encodes a protein of 147 amino acids with a calculated molecular weight of 16.1 kDa and a pI of 9.15.

Comparative Analysis of the Amino Acid Sequences of Triticale Phytocystatins

The amino acid sequences of the triticale phytocystatins contain four conserved regions of the characteristic domain, which determines the membership of the phytocystatin group (Fig. 1). The conserved regions were a glycine (G) (residue 49 in TrcC-1), the LARFAV sequence (residues 66–71 in TrcC-1), the QxVxG sequence (residues 93–97 in TrcC-1) and a tryptophan (W) (residue 124 in TrcC-1). In TrcC-5, a variation was observed within the LARFAV sequence with a glycine (G) at position 52 instead of an alanine (A) and with a tryptophan (W) at position 54 instead of a phenylalanine (F). The amino acid sequences of the triticale phytocystatins show diverse level of similarity to each other. TrcC-1 shares a 55% sequence identity with TrcC-4 and only a 23% sequence identity with TrcC-5. A low level of similarity was also observed between the TrcC-4 and TrcC-5 sequences (27%). The TrcCs amino acid sequence identity was also analysed in relation to other cereal plant phytocystatins. The highest level of identity was observed with wheat (from 84% to 99%) and barley (from 57% to 88%) amino acid
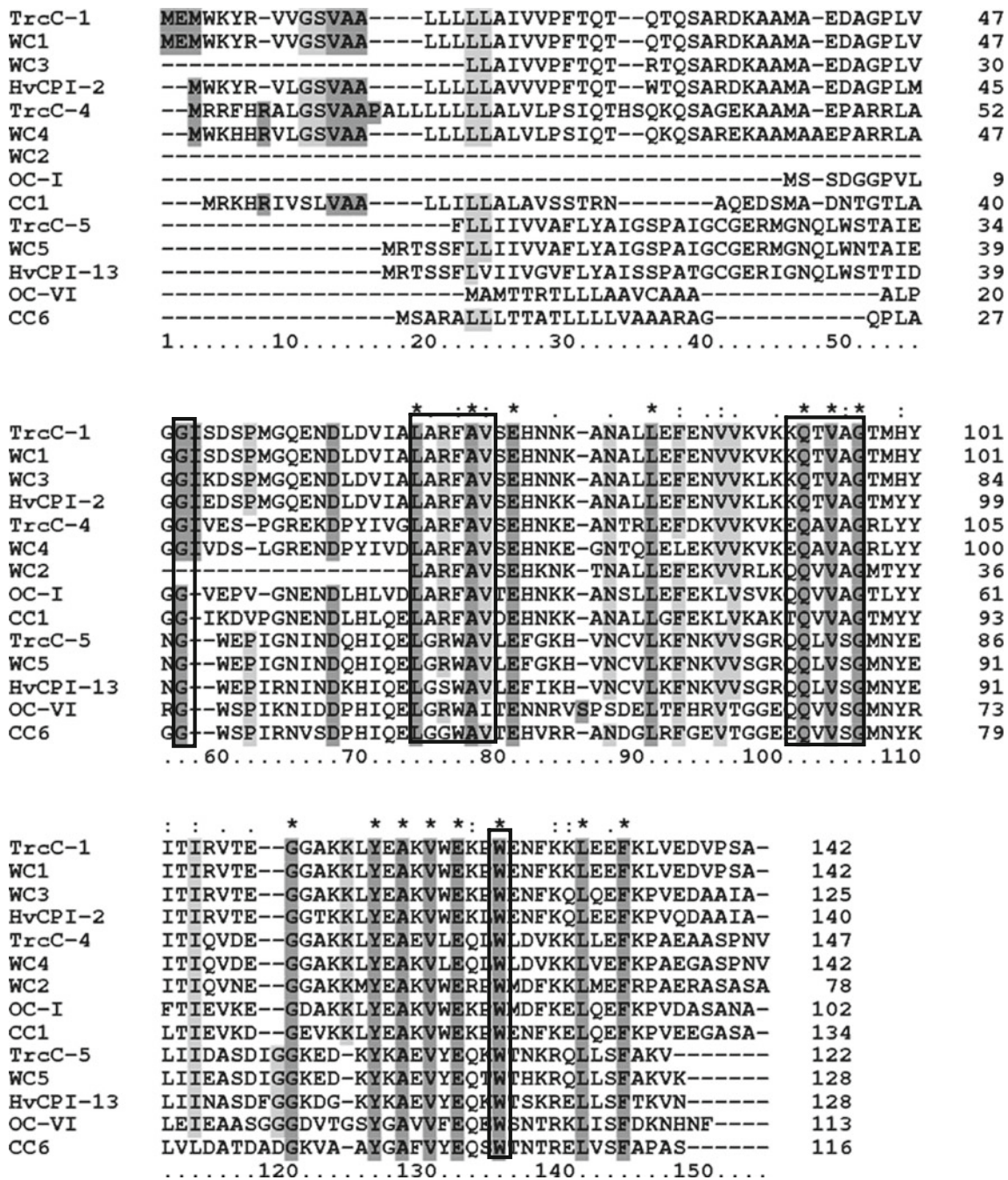
Fig. 1 Comparison of amino acid sequences of triticale phytocystatins with selected phytocystatins from the Poaceae family. CC corn phytocystatin, HvCPI barley phytocystatin, OC rice phytocystatin, TrcC triticale phytocystatin, WC wheat phytocystatin. The dark grey and asterisk marked amino acid residues indicate a 100% sequence identity, and the light grey marked residues indicate an 80% sequence identity. The conserved regions of phytocystatins are marked by frames

sequences, whereas the lowest was with phytocystatin sequences that were isolated from rice (from 41% to 64%) and corn (from 44% to 62%).

The phylogenetic analysis of the amino acid sequences of the triticale phytocystatins TrcC-1, 4 and 5, with five sequences of wheat (WC1–WC5), 12 sequences of barley (HvCPI-1–HvCPI-13), 11 sequences of rice (OC-I–OC-XII)

and 9 sequences of corn phytocystatins (CC1–CC10), was also performed (Fig. 2). The analysis showed that TrcCs were present in two of three dendogram groups that were identified by Martinez et al. (2009) (Fig. 2). The TrcC-1 and TrcC-4 phytocystatins were found in group A along with WC1–4 inhibitors from wheat, OC-I–III inhibitors from rice, CC1–5 inhibitors from corn and HvCPI-1–4
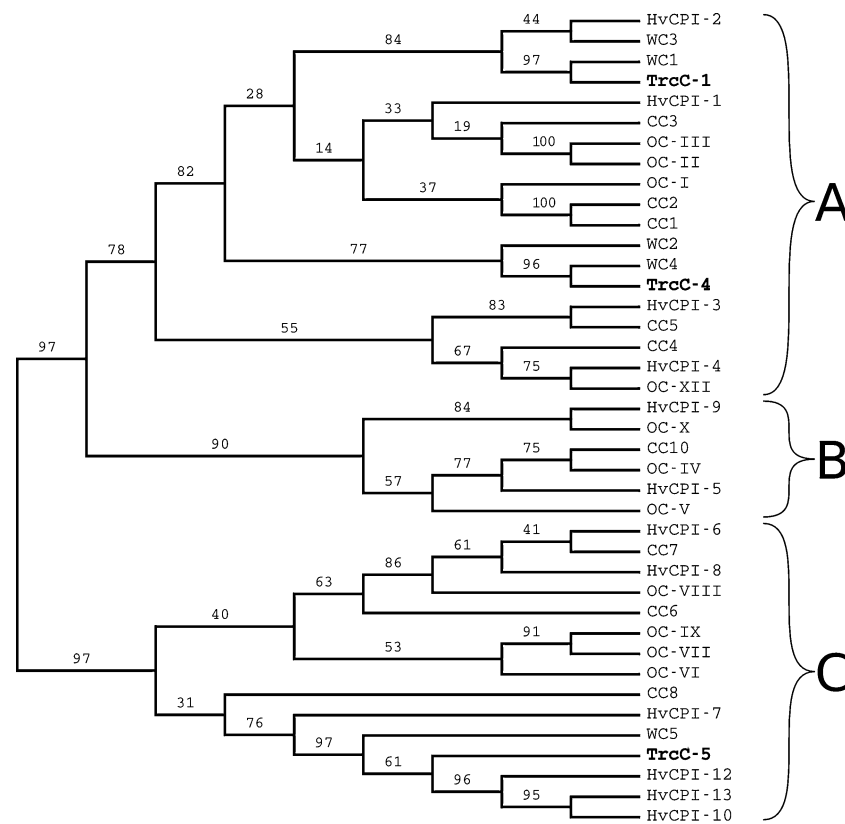
Fig. 2 The phylogenetic analysis of the amino acid sequences of triticale phytocystatins TrcC-1, 4, and 5 with 5 sequences of wheat (WC1–WC5), 12 sequences of barley (HvCPI-1–HvCPI-13), 11 sequences of rice (OC-I–OC-XII) and 9 sequences of corn phytocystatins (CC1–CC10). The phylogenetic tree was constructed using maximum likelihood method with JTT substitution evolution model and 100 bootstrap replicates on a sequence alignment produced with the ClustalW program. Numbers on branches are the maximum likelihood bootstrap support. Gene bank numbers corresponding to the sequences are as follows: TrcC-1 (GU395200), TrcC-4 (GU395201) and TrcC-5 (GU395202); HvCPI-1 (Y12068), HvCPI-2 (AJ748337), HvCPI-3 (AJ748338), HvCPI-4 (AJ748344), HvCPI-5 (AJ748340), HvCPI-6 (AJ748341), HvCPI-7 (AJ748345), HvCPI-8 (AJ748343), HvCPI-9 (AJ748339), HvCPI-10 (AJ748342), HvCPI-12 (AJ748347) and HvCPI-13 (AJ748348); for rice, OC-I (Os01g58890), OC-II (Os05g41460), OC-III (Os05g3880), OC-IV (Os01g68660), OC-V (Os01g68670), OC-VI (Os03g11180), OC-VII (Os03g11170), OC-VIII (Os03g31510), OC-IX (Os01g11160), OC-X (Os04g2250) and OC-XII (Os01g16430); for corn, CC1 (D63342), CC2 (D38130) and CC3–CC10 (BN000508–BN000514); for wheat, WC1 (AB038392), WC2 (AB038395), WC3 (AB038394), WC4 (AB038393) and WC5 (AF364099)

inhibitors from barley. TrcC-5 phytocystatin was located on a different branch of the phylogenetic tree (group C) than the other triticale phytocystatins, along with WC5 from wheat; OC-VI, VII and VIII from rice and CC6–8 from corn as well as HvCPI-6, 7, 8 and 10, 12, 13 from barley.

The Expression Pattern of *TrcCs* During the Development of Triticale Seeds

The mRNA levels of triticale phytocystatins during the seed development of the triticale were varied. However, an increasing trend in the level of gene expression during the first 2 weeks after pollination and a decrease in the level of mRNA thereafter, both in the Zorro and Disco cultivars (Fig. 3a, b), were common features to all three examined *TrcCs*. The expression of *TrcC-1* was detected during the whole period of seed development (Fig. 3a, b), with the highest level of mRNA observed between the seventh and tenth dap. Thereafter, the mRNA of *TrcC-1* significantly decreased. The changes in the mRNA levels of *TrcC-1* were similar for the Zorro and Disco cultivars. The expression of *TrcC-4* in seeds from both cultivars increased from seventh dap, reached a maximum at 13th dap, and then declined. However, in the Zorro cultivar, the *TrcC-4* transcript level was still significant at the 27th dap, while in the Disco cultivar, it was hardly noticeable already at the 20th dap. The mRNA of *TrcC-5* was observed over a significantly shorter period than the other phytocystatins and was visible only until the 17th dap in both examined triticale cultivars.

Expression Patterns of *TrcCs* During Germination of the Triticale Seeds

An increasing level of *TrcC-1* and *TrcC-4* mRNA was observed in the scutellum of seeds with the ongoing process of germination (Fig. 4a, b). In the case of TrcC-1, the

Fig. 3 The changes in mRNA levels of the triticale phytocystatins *TrcC-1*, *TrcC-4* and *TrcC-5* in developing seeds of the Zorro (**a**) and Disco (**b**) cultivars. RT-PCR analyses were performed with 20 ng of total RNA. The amplification of *18S rRNA* was used as the control to ensure that equal amounts of templates were added to each RT-PCR reaction. dap days after pollination
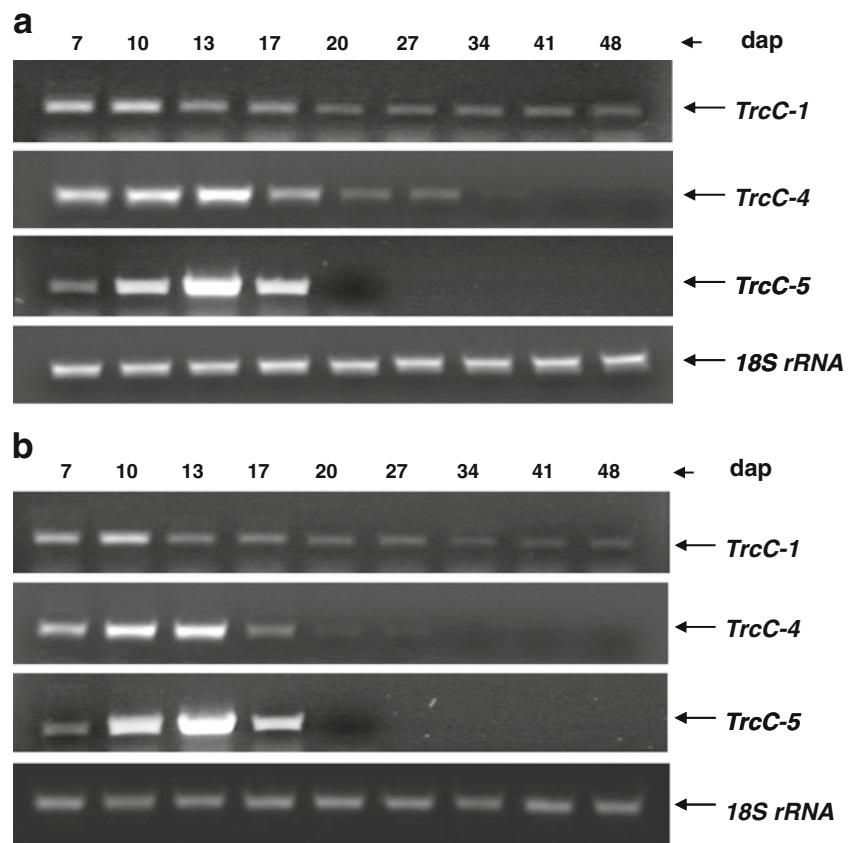
changes in the mRNA levels were similar for both examined triticale cultivars. The highest level of *TrcC-1* transcript was observed after 72 h of imbibition, and then, it

decreased. In case of *TrcC-4*, the mRNA levels varied between the examined cultivars. In the scutellum of the more resistant cultivar (Zorro), the level of *TrcC-4* mRNA

Fig. 4 The changes in the mRNA levels of the triticale phytocystatins, *TrcC-1* and *TrcC-4*, in the scutellum of the germinating triticale seeds in the Zorro (**a**) and Disco (**b**) cultivars. RT-PCR analyses were performed with 50 ng of total RNA. The amplification of *18S rRNA* was used as the control to ensure that equal amounts of templates were added to each RT-PCR reaction
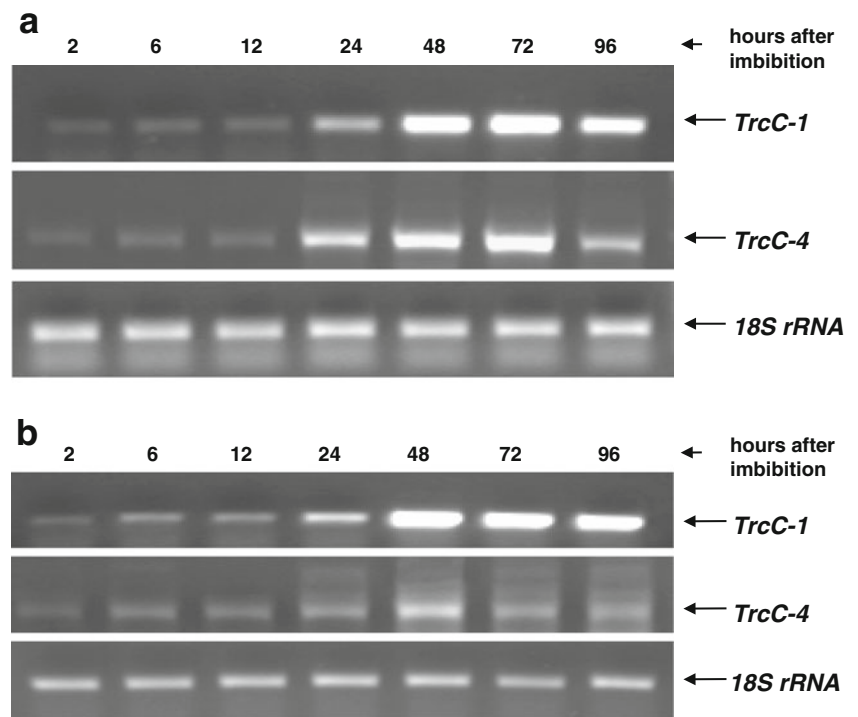
was its highest 72 h after imbibition, while in the scutellum of the more sensitive cultivar (Disco), the highest level of *TrcC-4* mRNA was observed already 48 h after imbibition. Moreover, in the Zorro cultivar, the level of *TrcC-4* was substantially higher than in the Disco cultivar.

In the endosperm of the germinating seeds, the level of *TrcC-1* mRNA was lower than in the scutellum, and it was constant in all examined stages of this process. The mRNA level of *TrcC-4* was low up to 48th h from imbibition, and thereafter, a significant increase of *TrcC-4* expression was observed both in the Disco and Zorro cultivars. After 96 h from imbibition, the *TrcC-4* transcript level substantially decreased again (Fig. 5a, b).

In the scutellum and endosperm of the germinating seeds, no expression of *TrcC-5* was observed.

Discussion

Three different nucleotide sequences of triticale phytocystatins were obtained and named *TrcC-1*, *TrcC-4* and *TrcC-5*. Within the deduced TrcC amino acid sequences, the presence of conserved regions characteristic for phytocystatins were observed: the LARFAV sequence, the QxVxG sequence, a glycine residue in the N-terminus and a tryptophan in the C-terminus, which determine the membership to the phytocystatin family and are vital for the activity of these inhibitors (Turk and Bode 1991). The conserved regions within the TrcC-1 and TrcC-4 sequences

were kept invariably, while some differences were observed within the LARFAV sequence of TrcC-5 (Fig. 1). Despite the observed changes, the essential residues for creating the complex with the enzyme were retained. In the research on the substitution of glutamine Q^{63} with leucine (L) in the sequence of $Q^{63}V^{64}V^{65}A^{66}G^{67}$ and arginine (R^{38}) with glycine (G) in the sequence of $L^{36}A^{37}R^{38}F^{39}A^{40}V^{41}$, Martinez et al. (2003) pointed to an important role of glutamine Q^{63} and arginine R^{38} in retaining the inhibitory properties of the HvCPI barley phytocystatins. The point mutations led to the lowering of the inhibitory activity of HvCPI and increased the inhibitory constant K_i (Martinez et al. 2003). However, according to the authors, certain differences within the conserved phytocystatin regions, which do not substantially influence the inhibitory activity, are allowed (Martinez et al. 2005). Apart from the determination of conserved regions, the fact that TrcCs belong to the phytocystatin family was confirmed by a high level of sequence identity with known phytocystatins from the Poaceae family. The highest percentage of the amino acid sequence identity was obtained with wheat inhibitors, from 84% (between WC4 and TrcC-4) to 99% (between WC1 and TrcC-1), and barley, from 57% (between HvCPI-2 and TrcC-4) to 88% (between HvCPI-2 and TrcC-1). Lower percentage of sequence identity was observed with the rice inhibitors, from 41% (between OC-VI and TrcC-5) to 63% (between OC-I and TrcC-1), and corn, from 44% (between CC6 and TrcC-5) to 62% (between CC1 and TrcC-1). The low level of identity between TrcC-5 and

Fig. 5 The changes in the mRNA levels of the triticale phytocystatins *TrcC-1* and *TrcC-4* in the endosperm of the germinating triticale seeds in the Zorro (**a**) and Disco (**b**) cultivars. RT-PCR analyses were performed with 50 ng of total RNA. The amplification of *18S rRNA* was used as the control to ensure that equal amounts of templates were added to each RT-PCR reaction
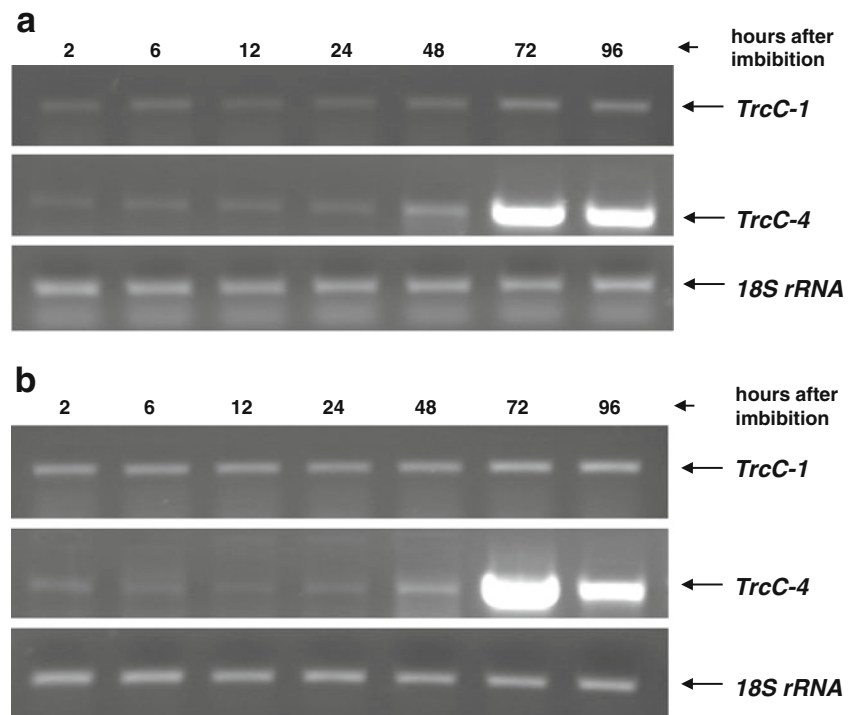
TrcC-1 or TrcC-5 and TrcC-4 sequences (23% or 27%, respectively) suggests that evolutionary processes are ongoing within this group of inhibitors. The aforementioned presumptions were confirmed by the phylogenetic analysis, according to which TrcC-5 is located on a different branch of the dendogram than TrcC-1 and TrcC-4 (Fig. 2). According to a classification proposed by Abraham et al. (2006) and Martinez et al. (2009), TrcC-5 is located in group C, of which representatives are involved in the regulation of endopeptidase activity in developing seeds. Among the members of this group, wheat (WC5) (Corre-Menguy et al. 2002), barley (HvCPI-6, 7, 8, 10, 12 and 13) (Martinez et al. 2009), rice (OC-VI, VII, VIII and IX) (Martinez et al. 2005) and corn phytocystatins (CC6, CC7 and CC8) (Massonneau et al. 2005) are present. Corre-Menguy et al. (2002) showed that the expression of the gene encoding WC5 occurs only during early stages of the seed development. The high level of mRNA and proteins of WC5 in the embryo and in the aleuron layer between 8 and 10 days of development indicates that this inhibitor might participate in the regulatory processes of embryogenesis and differentiation of the seed endosperm. Therefore, phytocystatin encoded by *TrcC-5*, which shares a 95% amino acid sequence identity with WC5 and which the highest level of mRNA was observed in the processes of seed formation and seed storage protein accumulation, might also participate in these processes by the protection of the storage proteins from the excessive activity of endogenous cysteine endopeptidases. However, the presence of WC5 protein until 20 days after pollination in the cells of the seed coat might point to an additional protective role of this inhibitor from exogenous proteinases of pests. *TrcC-5* gene expression was not observed in the scutellum and endosperms of the germinating seeds, which is additional evidence to suggest the fact that TrcC-5 is an inhibitor that might be connected specifically with the processes that take place during seed formation.

TrcC-1 and TrcC-4 belong to group A (Fig. 2) of the phytocystatins from the Poaceae family, among other already known inhibitors of wheat (WC1–4) (Kuroda et al. 2001), barley (HvCPI-1–4) (Abraham et al. 2006), rice (OC-I, II, III and XII) (Abe et al. 1987; Martinez et al. 2005) and corn (CC1–CC5) (Abe et al. 1994; Massonneau et al. 2005). According to the research performed by Abraham et al. (2006), the inhibitors from this group can regulate the activity of the endogenous cysteine endopeptidases both in vegetative and generative tissues and the control activity of exogenous enzymes from pests and pathogens. Despite the fact that TrcC-1 and Trc-4 belong to group A, they do not show a high level of sequence identity to each other (55%). A similar observation was reported for rice phytocystatins OC-I and OC-II present in the same group of inhibitors (Abe et al. 1987). Diversity of the rice

phytocystatins (OC-I and OC-II) was also observed in the gene expression profile and the ability to control the activity of various cysteine endopeptidases (Arai et al. 2002). While OC-I regulates the activity of endopeptidases from the papaine family whose expression was observed in dry (oryzain β) and germinating seeds (oryzain α and β), OC-II inhibits, similar to human cathepsin H, oryzain γ (Arai et al. 2002). The diversity within group A was also observed among wheat phytocystatins. Kuroda et al. (2001) detected mRNA of wheat phytocystatins from group A (WC1 – 4) in seed tissues in the first 2 weeks after pollination. The expression was again induced after sowing and was observed further in shoots and roots. However, in each case, the WC1 was expressed at a higher level than others. Moreover, immunostaining showed that only WC1 protein existed in the endosperm of seeds. Kuroda et al. (2001) suggest that WC1 controls several cysteine endopeptidases in all areas of wheat seeds, while WC4 probably has specified target enzymes. On the basis of differences observed in the sequences and expression profiles of *TrcC-1* and *TrcC-4*, it is possible that both triticale phytocystatins might perform various functions for triticale phytocystatins as the inhibitors of rice (OC-I and OC-II) and wheat (WC1–4). The highest level of *TrcC-1* mRNA was observed 10 days after pollination, after which the levels fell systematically (Fig. 3a, b). A low transcript level of this gene remained until the end of the process of seed formation in both the Zorro and Disco cultivars (Fig. 3a, b). These analyses suggest that TrcC-1, like some other members of group A, may control the activity of endogenous cysteine endopeptidases that participate in the processes ongoing during the whole period of seed development and maturation. The expression profile of *TrcC-1* mRNA in tissues of germinating seeds (Figs. 4 and 5) might also suggest the possible involvement of this inhibitor in the control of storage protein hydrolysis caused by the endopeptidases which are synthesised *de novo*. High-level mRNA of barley phytocystatin HvCPI-2, which shares an 88% sequence identity with TrcC-1, has been observed not only in the embryos of mature and germinating seeds (Martinez et al. 2005; Martinez et al. 2009), but the expression has also been detected in roots and leaves of barley (Martinez et al. 2009). Probably, phytocystatin HvCPI-2, like WC1, may protect the tissues of vegetative organs against pests and pathogens.

The increasing trend of gene expression, characteristic for the triticale phytocystatins, in the first 2 weeks of seed development and its gradual decline with the ongoing process have also been observed for *TrcC-4* (Fig. 3a, b). However, the gene expression of *TrcC-4* compared with *TrcC-1* was stronger in the first weeks after pollination, and it lasted for a much shorter period (Fig. 3a, b). The observations suggest that the control of the cysteine endopeptidase activities, which take part in the process of

the embryo development and the endosperm differentiation in the first weeks of seed development, could be the function of TrcC-4 activity, as was observed for WC4. The high level of TrcC-4 inhibitor between 10 and 13 days after pollination (Fig. 3a, b) might probably provide the protection of storage proteins synthesised in the third week of the seed formation. Differences in the dynamics and the levels of TrcC-4 mRNA, observed among the triticale cultivars, such as longer lasting and higher levels of mRNA in seeds of the Zorro cultivar when compared to the Disco cultivar, might suggest that TrcC-4 has influence on the increased resistance of the Zorro cultivar to preharvest sprouting.

The majority of the cysteine endopeptidases in seeds that participate in the degradation of seed storage proteins and are potentially involved in preharvest sprouting are synthesised *de novo* in the scutellum and aleuron layer of germinating cereal seeds (Prabucka and Bielawski 2004; Drzymała et al. 2009). The activity of these enzymes increases with the germination process. The expression of the genes encoding barley cysteine protease (EPB) in the first 24 h of imbibition was shown in the epithelium of the embryo, whereas from the second day of imbibition, the localisation of synthesis was in the aleuron layer of the starch endosperm (Mikkonen et al. 1996). High levels of mRNA of the wheat endopeptidases, triticain α and γ, which are homologous to rice oryzain α and γ, were shown in the embryo and in the aleuron layer in the second day of the germinating seeds. This level was maintained until the fifth day of the process (Kiyosaki et al. 2009). In the endosperm from triticale seeds, the presence of eight endopeptidases (EP1–EP8) hydrolyzing wheat storage protein, gliadin, was found on the third day after germination (Prabucka and Bielawski 2004). Therefore, the expression of TrcC-4, which was observed in the early hours after imbibition of triticale seeds, and its increasing expression until the third day of germination (Figs. 4a, b and 5a, b) might likely balance the increasing synthesis of cysteine endopeptidases in order to prevent the uncontrolled hydrolysis of seed storage proteins and results in preventing the preharvest sprouting in the ear.

References

Abe K, Emori Y, Kondo H, Suzuki K, Arai S (1987) Molecular cloning of a cysteine proteinase inhibitor of rice (oryzacystatin). Homology with animal cystatins and transient expression in the ripening process of rice seeds. J Biol Chem 262:16793–16797

Abe M, Abe K, Iwabuchi K, Domoto C, Arai S (1994) Corn cystatin I expressed in *Escherichia coli*: investigation of its inhibitory profile and occurrence in corn kernels. J Biochem 116:488–492

Abraham Z, Martinez M, Carbonero P, Diaz I (2006) Structural and functional diversity within the cystatin gene family of *Hordeum vulgare*. J Exp Bot 57:4245–4255

Arai S, Matsumoto I, Emori Y, Abe K (2002) Plant seed cystatins and their target enzymes of endogenous and exogenous origin. J Agric Food Chem 50:6612–6617

Bielawski W, Prabucka B (2001) Endopeptidases of triticale seeds. Biol Plant 44:283–288

Bielawski W, Dojczew D, Kączkowski J, Kolbuszewska-Podres W (1994) Enzymes of protein breakdown in germinating triticale grains resistant and susceptible to preharvest sprouting. Acta Physiol Plant 16:19–26

Bode W, Huber R (1992) Natural protein proteinase inhibitors and their interaction with proteinases. Eur J Biochem 204:433–451

Carrillo L, Martinez M, Ramessar K, Cambra I, Castanera P, Ortego F, Diaz I (2011) Expression of a barley cystatin gene in maize enhances resistance against phytophagous mites by altering their cysteine-proteases. Plant Cell Rep 30:101–112

Chomczynski P, Sacchi N (1987) Single-step method of total RNA isolation by a single extraction with an acid guanidinium thiocyanate-phenol-chloroform extraction. Anal Biochem 162:156–159

Corre-Menguy F, Cejudo FJ, Mazubert C, Vidal J, Lelandais-Briere C, Torres G, Rode A, Hartmann C (2002) Characterization of the expression of a wheat cystatin gene during caryopsis development. Plant Mol Biol 50:687–698

Dereeper A, Guignon V, Blanc G, Audic S, Buffet S, Chevenet F, Dufayard J-F, Guindon S, Lefort V, Lescot M, Claverie J-M, Gascuel O (2008) Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research Web Server Issue: W465-9.

Drzymała A, Prabucka B, Bielawski W (2009) Mobilizacja białek zapasowych ziarniaków zbóż. Post Biochem 55:447–455

Gaddour K, Vicente-Carbajosa J, Lara P, Isabel-Lamoneda I, Diaz I, Carbonero P (2001) A constitutive cystatin-encoding gene from barley (icy) responds differentially to abiotic stimuli. Plant Mol Biol 45:599–608

Gburek J, Gołąb K (1996) Cystatyny pochodzenia roślinnego. Post Biol Kom 23:385–398

Kiyosaki T, Asakura T, Matsumoto I, Tamura T, Terauchi K, Funaki J, Kuroda M, Misaka T, Abe K (2009) Wheat cysteine proteases triticain alpha, beta and gamma exhibit mutually distinct responses to gibberellin in germinating seeds. J Plant Physiol 166:101–106

Kuroda M, Kiyosaki T, Matsumoto I, Misaka T, Arai S, Abe K (2001) Molecular cloning, characterization and expression of wheat cystatins. Biosci Biotechnol Biochem 65:22–28

Margis R, Reis EM, Villeret V (1998) Structural and phylogenetic relationships among plant and animal cystatins. Arch Biochem Biophys 359:24–30

Martinez M, Lopez-Solanilla E, Rodriguez-Palenzuela P, Carbonero P, Diaz I (2003) Inhibition of plant-pathogenic fungi by the barley cystatin Hv-CPI (gene Icy) is not associated with its cysteine-proteinase inhibitory properties. Mol Plant Microbe Interact 16:876–883

Martinez M, Abraham Z, Carbonero P, Diaz I (2005) Comparative phylogenetic analysis of cystatin gene families from arabidopsis, rice and barley. Mol Genet Genomics 273:423–432

Martinez M, Cambra I, Carrillo L, Diaz-Mendoza M, Diaz I (2009) Characterization of the entire cystatin gene family in barley and their target cathepsin l-like cysteine-proteases, partners in the

hordein mobilization during seed germination. Plant Physiol 151:1531–1545

Massonneau A, Condamine P, Wiśniewski JP, Zivy M, Rogowsky PM (2005) Maize cystatins respond to developmental cues, cold stress and drought. Biochim Biophys Acta 1729:186–199

Mikkonen A, Porali I, Cercos M, Ho TH (1996) A major cysteine proteinase, EPB, in germinating barley seeds: structure of two intronless genes and regulation of expression. Plant Mol Biol 31:239–254

Mosolov VV, Valueva TA (2005) Proteinase inhibitors and their function in plant: a review. Biochem Microbiol 41:227–246

Muellenborn C, Krause J-H, Cerboncini C (2011) Analysis of differential transcript expression reveals time-dependent leaf responses to Sclerotinia sclerotiorum in wild and cultivated sunflower. Plant Mol Biol Rep 29:597–608

Prabu G, Kawar PG, Pagariya MC, Prasad DT (2011) Identification of water deficit stress upregulated genes in sugarcane. Plant Mol Biol Rep 29:291–304

Prabucka B, Bielawski W (2004) Purification and partial characteristic of major gliadin-degrading cysteine endopeptidase from germinating triticale seeds. Acta Physiol Plant 26:383–391

Shyu DJ, Chou WM, Yiu TJ, Lin CP, Tzen JT (2004) Cloning, functional expression, and characterization of cystatin in sesame seed. J Agric Food Chem 52:1350–1356

Travis J, Potempa J, Maeda H (1995) Are bacterial proteinases pathogenic factors? Trends Microbiol 3:405–407

Turk V, Bode W (1991) The cystatins: protein inhibitors of cysteine proteinases. FEBS Lett 285:213–219

Viswanathan A, Kuriakose B, Bharadwaj S, Thomas G (2011) Expression of aprotinin in anther causes male sterility in tobacco var Petit havana. Plant Mol Biol Rep 29:825–834

Volpicella M, Leoni C, Costanza A, De Leo F, Gallerani R, Ceci LR (2011) Cystatins, serpins and other families of protease inhibitors in plants. Curr Protein Pept Sci

Walsh TA, Strickland JA (1993) Proteolysis of the 85-kilodalton crystalline cysteine proteinase inhibitor from potato releases functional cystatin domains. Plant Physiol 103:1227–1234

Wu J, Haard NF (2000) Purification and characterization of a cystatin from the leaves of methyl jasmonate treated tomato plants. Comp Biochem Physiol C Toxicol Pharmacol 127:209–220

Zhang H, Hu Y, Wang C, Ji W (2011) Gene expression in wheat induced by inoculation with Puccinia striiformis west. Plant Mol Biol Rep 29:458–465